2011

Efficacy of segmental versus global core stabilization exercises for patients with chronic low back pain (LBP)

Hui En Gilpin
Pacific University

Efficacy of segmental versus global core stabilization exercises for patients with chronic low back pain (LBP)

Title: Efficacy of segmental versus global core stabilization exercises for patients with chronic low back pain (LBP).

Clinical scenario: At this facility, I have seen several patients with chronic LBP who have a core stabilization exercise program incorporated into their rehabilitation program. There are two physical therapists at the clinic, and one tends to start with segmental or local stabilization exercises involving learning to fire the deep core stabilizer muscles such as transversus abdominis (TA) and multifidus before progressing to global core stabilization. The other therapist usually goes straight to incorporating a global stabilization exercise program involving kinetic chains and muscular slings, and works on the larger core stabilizer muscles such as rectus abdominis, internal and external obliques, and erector spinae. I would like to know if there is a difference in outcomes between starting with segmental stabilization, or going straight to global stabilization without first teaching segmental stabilization for patients with chronic LBP not due to instability. Other interventions in a typical session include joint mobilizations and lower extremity strengthening exercises especially of the hip musculature.

Brief introduction: For the purposes of my clinical question, I want to know what the research says about the global and local or segmental core stabilization exercises for patients with chronic LBP not due to instability. The patients in the outpatient orthopedic facility I am working in often have LBP with weak core and hip musculature limiting functional mobility and activities. Between the two therapists working at this facility, one prefers to start with segmental stabilization and progress to global stabilization while the other therapist prefers to start working on global stabilization with this patient population. My question stems from these two different approaches I have observed in the clinic and whether one is more effective than the other in terms of reducing pain and improving function.

My clinical question: Is segmental stabilization more effective in reducing long term pain and improving long term functional outcomes than global stabilization for patients with chronic LBP?

Clinical question PICO:
Population: Adults aged 30-65 years with LBP not due to instability for at least 6 months affecting activities of daily living (ADLs), mobility, work or recreational activities
Intervention: Segmental core stabilization exercise program with conventional physical therapy (PT) including joint mobilizations, soft tissue mobilization (STM), stretching, modalities and/or education
Comparison: Global core stabilization exercise program with conventional PT as in the intervention group
Outcome: Pain relief as measured by the visual analogue scale (VAS) or Short-Form McGill Pain Questionnaire (SFMPQ) at discharge (up to 3 months from start of treatment) and at least 2 months after treatment has ended, Oswestry LBP disability questionnaire or Roland-Morris disability questionnaire (RMDQ) at discharge and at least 2 months after discharge

Overall clinical bottom line: Based on the results of the studies by Koumantakis *et. al* (2005) and Cairns *et. al* (2006), general exercise including global core stabilization with or without the addition of specific segmental core stabilization were both effective in decreasing pain as measured by the VAS and improving overall function as measured by the RMDQ in the long term (3 to 12 months after discharge). However, there is no statistical significance in results to show that one treatment is any more effective than the other in improving long term function and pain. This suggests that there is no added long term benefit to using segmental spinal stabilization exercises for patients with nonspecific chronic LBP. The main threats to internal validity in both studies are the lack of blinding of therapists and subjects, a large study loss of 30%, differences in VAS B scores (pain over last week) at baseline in the study by Koumantakis *et. al* (2005) and the lack of detailed protocol for general exercises in the conventional PT group in the study by Cairns *et. al* (2006). The first and last threats are minor, and the other two are moderate but were accounted for in the statistics with an intention to treat analysis and an ANCOVA. The main cost is time and financial cost of treatment, but the treatment is within reason to be covered by insurance. The results from both studies show that the addition of specific core stabilization exercises is not any more effective than conventional PT and general exercise including global core stabilization, but both are effective treatments in reducing pain and improving function. Thus, I would recommend that the clinician use discretion in determining whether to include specific segmental core stabilization in addition

to conventional treatment including global core stabilization for patients with nonspecific chronic LBP on a case by case basis.

Search terms: Low back pain, stabilization, outcome assessment, exercise therapy

Appraised by: Hui En Gilpin, SPT
School of Physical Therapy
College of Health Professions
Pacific University
Hillsboro, OR 97123
huiengilpin@pacificu.edu

Rationale for chosen articles: A systematic review by Rackwitz *et al.* (2006) was found on the topic of specific spinal stabilization exercises compared to general exercise programs or medications by a physician for populations with chronic LBP. None of the articles in the systematic review fit my clinical PICO well in terms of comparison group and outcomes, so I reviewed a list of other articles found using the search terms above. The systematic review also did not include some of the newer articles that have been published since 2004. After reviewing the list of articles, the following 3 articles were chosen because they were the best match to the clinical PICO in all aspects. The remaining articles were not chosen mainly because they did not have a matching comparison intervention or subject population as in my clinical PICO, or did not use an outcome measure I was looking for, and did not score any higher on the PEDro scale.

Article: Critchley DJ, Ratcliffe J, Noonan S, Jones RH, Hurley MV. Effectiveness and cost-effectiveness of three types of physiotherapy used to reduce chronic low back pain disability: a pragmatic randomized trial with economic evaluation. *Spine* 2007;32(14):1474-81.

PEDro score: 7/10
P: Patients aged 18 years or older with LBP for at least 12 weeks with or without leg symptoms or neurological signs, excluding those who had prior spinal surgery, physiotherapy for low back pain in the last 6 months, medical conditions such as rheumatological diseases or other disabilities rendering them unsuitable for group treatments of low back pain.
I: Spinal stabilization physiotherapy followed by group exercises that challenged spinal stability, for a maximum of 8 sessions each 90 minutes long
C: Individual physiotherapy not including specific spinal stabilization, for a maximum of 12 sessions each 30 minutes long or pain management program with group general exercises, for a maximum of 8 sessions each 90 minutes long
O: RMDQ, VAS, health-related quality of life measured with EQ-5D (EuroQoL) questionnaire, work participation measured by work days not working due to back pain in the previous 6 months, all measured at baseline, 6, 12 and 18 months

Article: Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Physical Therapy* 2005;85(3):209-25.

PEDro score: 6/10
P: Patients aged 25 to 50 years with history of recurrent LBP (repeated episodes in past year collectively lasting for less than 6 months) of nonspecific nature, and excluding those with prior spinal surgery, red flags, serious spinal pathology, nerve root pain signs or signs and symptoms of instability.
I: Specific trunk muscle stabilization exercise techniques combined with general exercise including global core stabilization and standardized patient education, for 45-60 minutes twice a week for 8 weeks
C: General exercise only including global core stabilization and patient education as in intervention group above, for 45-60 minutes twice a week for 8 weeks
O: RMDQ, pain using the SFMPQ and VAS, cognitive status using Pain Self-Efficacy Questionnaire, Tampa Scale of Kinesiophobia and Pain Locus of Control Scale, measured at baseline, 8 and 20 weeks

Article: Cairns MC, Foster NE, Wright C. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine* 2006;31(19):E670-81.

PEDro score: 7/10
P: Patients aged 18 to 60 years with recurrent LBP with or without radiating leg pain who had at least 1 previous episode of LBP that required alteration in normal activities or for which medical care or intervention was sought, and excluding those with psychologic distress, previous abdominal or spinal injury, presence of red flags, systemic illness, neurologic or muscular degenerative disorders or pregnancy or less than 1 year postpartum.
I: Segmental core stabilization with conventional PT as used within the UK clinical practice including manual and exercise (other than stability training) treatments and standardized educational information for a maximum of 12 sessions each 30 minutes long over 12 weeks
C: Conventional PT as in intervention group above for a maximum of 12 sessions each 30 minutes long over 12 weeks
O: RMDQ, pain using the SFMPQ and VAS, pscyhologic distress using the Modified Zung and Modified Somatic Perception Questionnaire Short-Form 36, measured at baseline, at discharge, 6 and 12 months following discharge

Table 1: Comparison of PEDro Scores

	Critchley DJ *et al.* (2007)	Koumantakis GA *et al.* (2005)	Cairns MC *et al.* (2006)
Random	1	1	1
Concealed allocation	1	1	1
Baseline comparability	1	0	1
Blind subjects	0	0	0
Blind therapists	0	0	0
Blind assessors	1	1	1
Adequate follow up	0	0	0
Intention to treat	1	1	1
Between group	1	1	1
Point estimates and variability	1	1	1
Total score	7/10	6/10	7/10

Table 1 shows a comparison of the PEDro scores. All 3 had similar PEDro scores, and were lacking in the same categories on the PEDro scale. None had blinding of subjects or therapists as it was not feasible since they could see what exercises they were doing. However, all 3 studies without blinding of subjects mentioned that subjects were not told anything specific about the two specific exercise programs in an effort to not bias them towards one treatment. All 3 articles also did not have adequate follow up, with greater than 15% study loss at the follow up assessments after termination of treatment. However, all 3 studies did include an intention to treat analysis to account for the large study loss. The study by Koumantakis *et al.* (2005) scored one lower on the PEDro scale than the other 2 articles because of baseline similarity between groups. Both groups were similar at baseline in all categories except for the VAS B (pain in the past week) scores. This is important because it is one of the outcomes of interest, and the raw data at different follow up times was not provided in the article, only change scores for each group.

Table 2: Other comparisons between articles

	Critchley DJ *et al.* (2007)	Koumantakis GA *et al.* (2005)	Cairns MC *et al.* (2006)
Power analysis	90% power based on 12 points on RMDQ	80% power based on 2.5 points on RMDQ	90% power based on 5 points on RMDQ
Population	Similar	Similar	Similar
Intervention	Local core stabilization	Local then global core stabilization	Local core stabilization with conventional PT
Comparison	Conventional PT Pain class	Global core stabilization	Conventional PT
Outcomes	RMDQ, SFMPQ, VAS	RMDQ, SFMPQ, VAS B	RMDQ, SFMPQ, VAS
Protocol	Protocol not provided	Protocol provided	Protocol provided

All 3 articles performed a power analysis to determine a priori sample size, using the criteria summarized in Table 2. Also, only the articles by Koumantakis *et al.* (2005) and Cairns *et al.* (2006) included an appendix with more specifics on the protocol for the exercise programs. All 3 articles had similar populations and outcome measures as in the clinical PICO, but the article by Koumantakis *et al.* (2005) had intervention and comparison groups most similar to the clinical PICO. Only the protocol for the specific spinal stabilization was provided in the article by Cairns *et al.* (2006), and it is unclear what exercises in the conventional group was done and whether global core stabilization was included. There was mention that exercises in the comparison group included specific trunk muscle retraining and general spinal mobility in the study by Critchley *et al.* (2007), but without details on the protocol, it is unclear whether global core stabilization was included.

I decided to put more weight on the similarity in the article PICO to my clinical PICO, hence I chose the article by Koumantakis *et al.* (2005) over the other two even though it scored the lowest on the PEDro (but only by 1 point) because of the closest similarity to my clinical PICO. Both articles by Critchley *et al.* (2007) and Cairns *et al.* (2006) scored the same on the PEDro and were lacking in the same areas, had large study losses and were done in the United Kingdom (UK). However, the article by Cairns *et al.* (2006) was a multicenter study, did have a protocol provided even though only for the specific spinal stabilization group, and the outcome measures reported included mean differences within and between groups, with reliability and minimal clinically important differences (MCIDs) also provided. The article by Critchley *et al.* (2007) was a single center study, did not have any protocol provided, and the outcome measures are provided in a way that mean differences cannot be calculated, with no mention of reliability or MCID for the outcome measures. Thus, based on the above comparisons, I have chosen to write this critically appraised paper on the articles by Koumantakis GA *et al.* (2005) and Cairns MC *et al*. (2006).

Article: Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Physical Therapy* 2005;85(3):209-25.

Clinical bottom line: The results of this study suggest both specific or segmental core stabilization and general exercise including global core stabilization were effective in decreasing pain as measured by the VAS B (pain over last week) and improving overall function as measured by the RMDQ. The differences within groups were both clinically important at the 3 month follow up, but had large 95% confidence intervals (CI). The results did not show a statistically significant or clinically important difference between groups for VAS B at discharge and the 3 month follow up. There was a statistical significance between groups in favor of the general exercise group for RMDQ at discharge with a minor clinically important difference (greater than the lower limit of the MCID but with a large 95% CI), but the effects were no longer statistically significant at the 3 month follow up. This suggests that there is no short term or added long term benefit to using segmental spinal stabilization exercises for patients with nonspecific chronic LBP. The primary threats to internal validity are the lack of blinding of therapists and subjects, a 30% study loss and differences in VAS B scores at baseline. The first threat is minor, and the last two are moderate threats. The large study loss was taken into account by the authors with an intention to treat analysis and using a large enough sample size determined by a power analysis. The differences at baseline in VAS B scores could signify that randomization was not completely successful, but the authors accounted for this by adjusting the VAS B scores for differences at baseline using an ANCOVA. Based

on this study, the addition of specific core stabilization exercises is not any more effective than general exercise including global core stabilization, but both are equally effective treatments in reducing pain and improving function, thus it should be left up to the clinician's discretion of whether to add in segmental core stabilization or not for patients with nonspecific chronic LBP on a case by case basis.

Article PICO:
Population: Patients aged 25 to 50 years with history of recurrent LBP (repeated episodes in past year collectively lasting for less than 6 months) of nonspecific nature, and excluding those with prior spinal surgery, red flags, serious spinal pathology, nerve root pain signs or signs and symptoms of instability.
Intervention: Specific trunk muscle stabilization exercise techniques combined with general exercise including global core stabilization and standardized patient education, for 45-60 minutes twice a week for 8 weeks
Comparison: General exercise only including global core stabilization and patient education as in intervention group above, for 45-60 minutes twice a week for 8 weeks
Outcome: RMDQ, pain using the SFMPQ and VAS, cognitive status using Pain Self-Efficacy Questionnaire, Tampa Scale of Kinesiophobia and Pain Locus of Control Scale, measured at baseline, 8 and 20 weeks

Blinding: The single assessor was blinded to group allocation. Both subjects and therapists could not be blinded to group allocation as they could see which exercise group they were in, but the authors reported that subjects were not aware of the theoretical bases of each of the exercise regimens in order to not bias them towards one treatment.

Controls: The group receiving general exercise including global core stabilization served as the control group (CG).

Randomization: Subjects were randomized into 2 groups, a treatment group (TG) receiving specific core stabilization (n=29) and a CG receiving general exercise only (n=26). Randomization was done via a computer generated random number sequence. Randomization was concealed by using an independent trial manager not involved in the study and keeping the randomization codes in a sealed envelope with consecutive numbering. Randomization appears successful. At baseline there were no significant differences between groups for age, gender, body weight, height, body mass index, history of LBP, SFMPQ, VAS C (pain in past month) and RMDQ scores. Only the VAS B (pain in past week) scores at baseline were different between groups at baseline.

Study: Fifty-five subjects were recruited from the local orthopedic clinic of a local hospital and several general practitioner's practices. Inclusion criteria were that subjects had a history of recurrent LBP (repeated episodes over the last year collectively lasting for less than 6 months) of a nonspecific nature, were medically fit with no heart problems, pregnancy or inflammatory arthritis, willing to participate in the exercise program and able to travel independently to the hospital. Exclusion criteria were prior spinal surgery, presence of "red flags" and signs and symptoms of instability. The subjects were randomized into 2 groups, a TG receiving specific core stabilization exercises and a CG receiving general exercise only. Both groups included a warm-up period consisting of stretches and stationary bike for 10-15 minutes. The TG received individual training for firing transversus abdominis (TA) and lumbar multifidus in multiple positions, with increase in holding time and number of contractions as the subjects progressed over time. Palpation on these deep stabilizers by the therapists was done to ensure correct muscle firing. Once subjects had achieved the specific spinal stabilization in low load positions, it was incorporated into light functional tasks involving spinal or limb movements, and finally integrated into heavier load functional tasks with exercises similar to those performed by subjects in the CG. In the CG, subjects performed general exercise only, consisting of global core stabilization with exercises activating the extensor (paraspinals) and flexor (abdominal) muscle groups. Both groups received treatment for 45-60 minutes each session, 2 times a week for 8 weeks. However, the total exercise time in the CG was kept at half that of the TG in an effort to balance the groups with respect to estimated total force output of the trunk muscles targeted by the exercises. The first session was on an individual basis, with subsequent exercises done in groups of 5-7 subjects in the TG and up to 10 subjects in the CG. Subjects were also

asked to repeat the exercises at home for a maximum of half hour 3 times a week. Subjects in both groups also received an information booklet for pain education.

Outcome measures: The outcome measures of interest are the RMDQ and VAS B (pain in past week) scores. These outcome measures were recorded at baseline, discharge (8 weeks) and at the 3 month follow up (20 weeks). The VAS score was recorded on a 0 to 100 mm scale. The authors looked at VAS A for current pain intensity, VAS B for average pain over the past week and VAS C for average pain over the past month and determined that the VAS B was the most reliable (ICC=0.88) then VAS C (ICC=0.77), with VAS A as the least reliable (ICC=0.46). Thus only VAS B and C scores were reported by the authors, and only VAS B will be considered here as it was the most reliable score. The authors did mention that the RMDQ has clinically acceptable reliability, validity and responsiveness. The authors did not discuss the minimal clinically important difference (MCID) for the VAS B score for this population, but mentioned an MCID of 2.5-4 points on the RMDQ. An MCID value of 10-20 mm on the VAS B scale has been identified for the population with LBP (Beurskens *et. al*, 1996, Ostelo *et. al*, 2005).

Study losses: There was a study loss of 10 out of 55 subjects (18% study loss) at discharge, with 5 subjects from each group. Most of the dropouts were due to time constraints, and 2 subjects from the TG were due to increased pain during the exercise program. At the 3 month follow up, there was an additional loss of 7 subjects (3 from the TG and 4 from the CG) because they did not return their questionnaires for unknown reasons, resulting in a total study loss of 17 out of 55 subjects (31%). The authors did mention performing an intention to treat analysis alongside a per protocol analysis, comparing the two results. The missing data for intention to treat analysis were handled in a relatively conservative approach by using group means in place of the missing values. The authors reported that results were the same for all outcome measures with both the intention to treat and per protocol analyses.

Summary of internal validity: Randomization of the subjects, similarity of subjects at baseline except for VAS B scores, blinding of assessors to group allocation, detailed protocol provided, and similarity in treatment time and administration are strengths of internal validity for this study. The study did mention a power analysis was done assuming 80% power and 5% significance to detect an MCID of 2.5 points on the RMDQ between groups, and a sample size of 28 subjects per group was required, or total of 56 subjects for the two groups in this case. The primary threats to internal validity include the lack of blinding of therapists and subjects, large study losses and differences in VAS B scores at baseline. The lack of blinding of therapists and subjects is a minor threat since the one therapist teaching exercises to both groups still treated each subject individually according to standard practice and best clinical judgment and subjects were not told whether one treatment was better than another. There was a large study loss of 10 out of 55 subjects or 18% study loss at discharge and 17 out of 55 subjects or 31% study loss at the 3 month follow up. This is a moderate threat, but an intention to treat analysis was done to account for the large study loss. The difference in baseline VAS B scores in a moderate threat as that means there could have been some differences between groups and randomization was not completely successful, but the VAS B scores were adjusted for the differences in baseline using an ANCOVA. Overall, there are no major threats to internal validity.

Evidence: The authors reported statistically significant differences in RMDQ and VAS B scores within both the TG and CG at discharge and the 3 month follow up, and between groups for the RMDQ score at discharge in favor of the CG, but not between groups at discharge for the VAS B score or at the 3 month follow up for RMDQ and VAS B scores. All the outcome measures between groups were analyzed using parametric statistical tests for statistical significance with VAS B data adjusted for differences in baseline using an ANCOVA. The authors reported that the per protocol and intention to treat analyses had similar results and only the numbers from the intention to treat analysis are provided. Tables 3, 4 and 5 summarize the data from the article.

Table 3: RMDQ and VAS B mean scores with standard deviation (SD)

	TG (n=29)		CG (n=26)	
	RMDQ (0-24)	VAS B (100mm)	RMDQ (0-24)	VAS B (100mm)
Baseline	9.2±4.6	26.9±20.6	11.3±5.2	40.2±24.6
Discharge (8 weeks)	5.1±4.0	12.3±13.7	4.7±3.5	21.3±17.3
3 month follow up (20 weeks)	4.5±3.8	15.8±15.3	5.2±3.5	17.8±14.2

Table 4: Within group mean differences for RMDQ and VAS B scores with SD and 95% CI

	TG (n=29)		CG (n=26)	
	RMDQ (0-24)	VAS B (100mm)	RMDQ (0-24)	VAS B (100 mm)
Baseline to discharge (8 weeks)	-4.05±3.26#* (-7.31 to -0.79)	-18.18±18.80#* (-36.98 to 0.62)	-6.60±4.97#* (-11.57 to -1.63)	-14.92±16.52#* (-31.44 to 1.60)
Baseline to 3 month follow up	-4.65±3.26#* (-7.91 to -1.39)	-15.16±19.10#* (-34.26 to 3.94)	-6.03±4.98#* (-11.01 to -1.05)	-17.78±19.70#* (-37.48 to 1.92)

# indicates a statistically significant difference was found
* indicates the mean difference (but not 95% CI) exceeded MCID

Table 5: Between groups (TG to CG) mean difference for RMDQ and VAS B change scores with 95% CI

	RMDQ (0-24)	VAS B (100mm)
Baseline to discharge (8 weeks)	2.55 (0.30 to 4.81)#*	-3.26 (-10.15 to 3.63)
Baseline to 3 month follow up	1.38 (-0.87 to 3.64)	2.62 (-4.58 to 9.82)

* indicates a statistically significant difference was found
* indicates the mean difference (but not 95% CI) exceeded MCID

The authors provided the mean differences within and between groups with intention to treat analysis. Since raw data was not provided at discharge and at the 3 month follow up, all effect sizes could not be calculated. There were no statistically significant differences between the intention to treat analysis and per protocol analysis. There was a clinically important difference within groups for RMDQ and VAS B scores, but with large 95% CI where the lower limits of the 95% CI fell below the MCID (2.5-4 points for RMDQ and 10-20 mm for VAS B) in all cases. There was also a change from negative to positive on the lower limit of the 95% CI for VAS B scores, which means that if repeated enough times, the pain as measured by the VAS B score could have had the opposite effect and increased within groups. No statistically significant differences and hence no clinically important differences were found between groups for VAS B scores from baseline to discharge and the 3 month follow up, and the RMDQ scores at the 3 month follow up. A statistically significant difference between groups in favor of the CG for RMDQ score at discharge was found, but was only barely greater than the lower end of the MCID of 2.5 points. However, the 95% CI is large and the lower limit falls well below the MCID. The results suggest that there could be a clinically important decrease in RMDQ and VAS B scores and hence improvement in function and LBP relief that is maintained up to 3 months after treatment using either specific or global stabilization exercise regimens. However, without statistical significance between groups for VAS B scores at discharge and the 3 month follow up and RMDQ scores at the 3 month follow up, and minor clinically important difference between groups for RMDQ scores at discharge with a large 95% CI, there is no strong evidence to suggest that the addition of specific spinal stabilization exercises including TA and lumbar multifidi is any more effective than general exercises including global core stabilization in pain reduction and improvement in function for those with chronic LBP.

Applicability of study results:

Similarity to my patients: The subjects in this study are similar to my patient population in terms of inclusion and exclusion criteria, and has a similar age range but with a slightly younger population (25 to 50 years compared to 30 to 65 years in the clinical PICO).

Benefits vs. costs: The main cost is the time spent and the financial cost of physical therapy treatment for 16 sessions each 45-60 minutes long, but was similar between groups. The possible benefit is long term improvement in function and pain relief up to 3 months after discharge, but both groups received the same benefit in the long term.

Feasibility of treatment: Both the specific and global core stabilization exercise programs are easily provided by a physical therapist, and are already commonly provided in physical therapy treatments for chronic LBP. The financial cost of 16 sessions of physical therapy treatment for 45-60 minutes each is also within reason to be covered by insurance, but is on the upper limit of typical number of visits. The protocol on the specific spinal stabilization exercises, general exercises including global core stabilization and pain education was described sufficiently in the article and can easily be reproduced in a clinical setting.

Summary of external validity: The subject sample is similar to the clinical population of interest, except for the slightly younger age range of subjects in the study. A minor threat to external validity is that subjects were sampled from only one local orthopedic center. However, the authors accounted for threats to external validity by having an adequate number of subjects. The authors had performed a power analysis and determined the a priori sample size needed to control for threats to external validity, and thus results may be more easily generalized to a wider population range.

Article: Cairns MC, Foster NE, Wright C. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine* 2006;31(19):E670-81.

Clinical bottom line: The results of this study suggest both segmental core stabilization with conventional PT and conventional PT including general exercise and manual therapy were effective in decreasing pain as measured by the VAS and improving overall function as measured by the RMDQ. The differences within groups were both statistically significant and clinically important at the 12 month follow up. The results did not show a statistically significant or clinically important difference between groups, suggesting that there is no added benefit to using segmental spinal stabilization exercises for patients with nonspecific chronic LBP. The primary threats to internal validity are the lack of blinding of therapists and subjects, the lack of detailed protocol for the general exercises in the conventional PT group and a 30% study loss. The first two threats are minor, and the third is a moderate threat, however was taken into account by the authors with an intention to treat analysis and using a large enough sample size determined by a power analysis. Based on this study, the addition of specific core stabilization exercises to conventional PT consisting of general exercise and manual therapy is not any more effective than just conventional PT alone, but both are effective treatments in reducing pain and improving function, thus it should be left up to the clinician's discretion of whether to add in segmental core stabilization or not for patients with nonspecific chronic LBP on a case by case basis.

Article PICO:
Population: Patients aged 18 to 60 years with recurrent LBP with or without radiating leg pain who had at least 1 previous episode of LBP that required alteration in normal activities or for which medical care or intervention was sought, and excluding those with psychologic distress, previous abdominal or spinal injury, presence of red flags, systemic illness, neurologic or muscular degenerative disorders or pregnancy or less than 1 year postpartum.
Intervention: Segmental core stabilization with conventional PT as used within the UK clinical practice including manual and exercise (other than stability training) treatments and standardized educational information for a maximum of 12 sessions each 30 minutes long over 12 weeks
Comparison: Conventional PT as in intervention group above for a maximum of 12 sessions each 30 minutes long over 12 weeks
Outcome: RMDQ, Oswestry Disability Index, pain using the SFMPQ and VAS, pscyhologic distress using the Modified Zung and Modified Somatic Perception Questionnaire (MSPQ), Short-Form 36 (SF-36), measured at baseline, at discharge, 6 and 12 months following discharge

Blinding: It is not possible to have subjects and therapists blinded in this case as they could see which exercise group they are in, but the authors reported that subjects were naïve to allocation in order not to bias them towards one treatment. Therapists had no influence over randomization and group allocation, and since the outcomes consisted of patient-completed measures only, it would be within reason to assume that the assessor can be considered blinded to allocation even though the study does not specifically state whether the researcher analyzing the data was blinded to group allocation.

Controls: The group receiving conventional treatment consisting of manual treatment and exercises excluding low load, high repetition muscle activity used in specific spinal stabilization as taught in the treatment group (TG) served as the control group (CG).

Randomization: Subjects were randomized into the TG receiving specific spinal stabilization exercises (n=47) and the CG receiving conventional treatment (n=50). Randomization was achieved using an adaptive stratified randomization procedure, assessing the participants' characteristics against 3 categories. An independent observer determined patient allocation with a coin flip, thus keeping randomization concealed. Randomization was successful, with no significant differences in age, height, weight, gender, duration of episode, RMDQ, Oswestry disability index, VAS, SF-36 and MSPQ scores at baseline.

Study: Ninety-seven subjects aged 18 to 60 years were recruited from several centers based in secondary care in the United Kingdom (UK). Inclusion criteria were LBP with or without radiating leg pain and with a minimum of 1 episode of LBP resulting in alteration of normal activities or for which medical intervention had been sought, RMDQ score ≥5 and sufficient proficiency in English to complete the self-report questionnaires. Exclusion criteria included red flags, abdominal or spinal surgery in the last 12 months, systemic illness, neurologic or muscular degenerative disorders, pregnancy or less than 1 year postpartum, or evidence of psychologic distress. The 97 subjects were randomized into 2 groups. The TG (n=47) received endurance training for the deep abdominal (TA) and back extensor (lumbar multifidi) muscles, with diagnostic ultrasound available to ensure that the deep core stabilizers were being fired. The details of the specific spinal stabilization regime used are outlined in the appendix provided in the study. The CG (n=50) received conventional treatment consisting of exercises excluding low load, high repetition muscle activity used in specific spinal stabilization as taught in the TG. Both groups also received standardized educational information in a booklet, as well as other manual and exercise treatments based on current UK clinical practice, but hydrotherapy, back school or other group therapy was prohibited. Treatment was individualized at the discretion of the clinician. Both groups received treatment up to a maximum of 12 treatment sessions over 12 weeks, with the initial assessment for 60 minutes and follow up sessions of 30 minutes each.

Outcome measures: The outcome measures of interest here are the RMDQ and VAS scores, which were recorded at baseline, discharge, and at the 6 and 12 month follow ups. All outcomes were done via patient-completed questionnaires. The authors mentioned that the RMDQ is advocated as a "core" outcome measure in LBP trials, has proven reliability and validity, and known test-retest reliability in various settings. There was no mention of reliability or validity of the VAS score in this study. The authors did discuss the MCID for the RMDQ and VAS scores for this population, reporting an MCID value of 3-5 points on the RMDQ and 1-1.8cm on the 10cm VAS scale for the population with LBP.

Study losses: At discharge, there was a study loss of 17 out of 97 subjects (18% study loss), 6 from the TG and 11 from the CG. There was a study loss of 29 out of 97 subjects (30% study loss) at the 12 month follow up, 14 from the TG and 15 from the CG. Losses were from subjects who did not complete a course of treatment or who failed to respond to the 12 month follow up. An intention to treat analysis was carried out in this study to address potential biases caused by incomplete follow up using the "best-case" and "worst-case" scenarios.

Summary of internal validity: Randomization of the subjects, similarity of subjects at baseline, large sample size, blinding of assessors to group allocation, detailed protocol for specific spinal stabilization group and pain education provided, similarity in treatment time and administration are strengths of internal validity for this study. The study did mention a power analysis was done assuming 90% power and 5% significance to detect an MCID of 5 points on the RMDQ between groups from baseline to the 12 month follow up, and a total sample size of 92 was required. The total sample size was based on a standard deviation (SD) of 6 points identified from pilot work, and allowed 10% attrition at each follow up point. The primary threats to internal validity include the lack of blinding of therapists and subjects, study losses and a lack details on the protocol for the exercises taught in the CG. The lack of blinding of therapists and subjects is a minor threat since the therapists still treated each subject individually according to standard practice and best clinical judgment and subjects were not told whether one

treatment was better than another. There was a large study loss of 17 out of 97 subjects or 18% study loss at discharge and 29 out of 97 subjects or 30% study loss at the 12 month follow up. This is a moderate threat, but an intention to treat analysis was done to account for the large study loss. The lack of protocol for the exercises in the CG is a minor threat because it makes it more difficult to reproduce the study and to know how much global core stabilization was included. Overall, there are no major threats to internal validity.

Evidence: The authors reported statistically and clinically significant differences between baseline and 12 month follow up RMDQ and VAS scores within both the TG and the CG groups, but neither statistically nor clinically significant differences between groups with and without intention to treat analysis. All of the outcome measures between groups were analyzed using ANOVA for statistical significance. The outcome measures I am interested in are the RMDQ and VAS scores at discharge and at the 12 month follow up. Only the baseline scores and change scores from baseline to the 12 month follow up are provided, and the authors mentioned that broadly, the findings from the 12 month follow up were similar to those seen at discharge and the 6 months follow up, and did not include the numerical data for discharge and the 6 months follow up. Tables 6, 7 and 8 summarize the data from the article.

Table 6: RMDQ and VAS baseline scores with standard deviation (SD)

	TG (n=47)		CG (n=50)	
	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)
Baseline score	10.4±4.3	4.2±2.0	10.3±4.1	4.22±2.3

Table 7: Mean difference for RMDQ and VAS scores within and between groups with 95% confidence interval (CI) with intention to treat analysis

	TG (n=47)		CG (n=50)		Between groups	
	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)
Baseline to 12 month follow up	-5.1#+ (-6.3, -3.9)	-1.5#* (-2.1, -0.9)	-5.4#+ (-6.5, -4.2)	-1.9#* (-2.5, -1.3)	-0.4 (-2.0, -1.3)	0.4 (-1.2, 0.5)

# indicates a statistically significant difference was found
+ indicates the mean difference and 95% CI exceeded MCID
* indicates the mean difference (but not 95% CI) exceeded MCID

Table 8: Mean difference for RMDQ and VAS scores within and between groups with 95% CI per protocol

	TG (n=33)		CG (n=35)		Between groups	
	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)
Baseline to 12 month follow up	-4.5#* (-6.2, -2.9)	-1.2#* (-2.1, -0.4)	-5.2#+ (-6.7, -3.6)	-1.8#* (-2.6, -0.9)	0.6 (-2.9, -1.7)	-0.5 (-1.75, 0.72)

# indicates a statistically significant difference was found
+ indicates the mean difference and 95% CI exceeded MCID
* indicates the mean difference (but not 95% CI) exceeded MCID

The authors provided the mean differences within and between groups with and without intention to treat analysis. Since raw data was not provided at discharge and at the 12 month follow up, all effect sizes and mean differences at discharge could not be calculated. There were no statistically significant differences between the intention to treat analysis and per protocol analysis. There was a clinically important difference within groups for RMDQ scores (greater than MCID of 3-5 points) taking into account the 95% CI, and a clinically important difference within groups for VAS scores, but the lower limit of the 95% CI fell below the MCID (1-1.8 cm). No statistically significant differences and hence no clinically important differences were found between groups for RMDQ or VAS scores from baseline to the 12 month follow up. The results suggest that there could be a clinically important decrease in RMDQ and VAS scores and hence improvement in function and LBP relief that is maintained up to 12 months after treatment using either treatment protocol. However, there is no statistically significant difference in RMDQ and VAS scores between the spinal stabilization exercise and conventional treatment groups, which suggest that the addition of specific spinal stabilization exercises including TA and lumbar multifidi is not any more

effective than general exercises in conventional treatment in pain reduction and improvement in function for those with chronic LBP.

Applicability of study results:
Similarity to my patients: The subjects in this study are similar to my patient population in terms of inclusion and exclusion criteria, and had similar age range except for younger patients were included (18 to 60 years in the study compared to 30 to 65 years in the clinical PICO). However, they were sampled from the UK, and there could be a cultural difference affecting the rating of function on the RMDQ and pain on the VAS.

Benefits vs. costs: The mean number of treatment sessions for the conventional treatment group was 5.9 (SD 2.3) and 7.5 (SD 2.5) for the specific stabilization group. The main cost is the time spent and the financial cost of physical therapy treatment for up to 12 sessions each 30 minutes long, but was similar between groups, with the specific stabilization group slightly more costly due to greater number of treatment sessions. The benefit is long term improvement in function and pain relief up to 12 months after discharge, but both groups received the same benefit.

Feasibility of treatment: Both the specific stabilization and conventional treatments are easily provided by a physical therapist, and are already commonly provided in physical therapy treatments for chronic LBP. The financial cost of up to 12 sessions of physical therapy treatment for 30 minutes each is also within reason to be covered by insurance. The protocol on the specific spinal stabilization exercises and pain education was described sufficiently in the article and can easily be reproduced in a clinical setting.

Summary of external validity: The subject sample is similar to the clinical population of interest, with the exception of the slightly younger age range of subjects in the study and culture differences that may affect disability and pain perception and thus rating of function on the RMDQ and pain on the VAS scale. Another strength is that this was a multicenter study, with subjects sampled from more than one facility, and thus results may be more easily generalized to a wider population range. Also, a power analysis was performed to determine the a priori sample size needed to control for threats to external validity.

Synthesis/Discussion:

The purpose of this paper was to determine if specific local or segmental core stabilization was more effective than global core stabilization in improving functional outcomes and pain relief in patients with chronic LBP. Both studies reviewed had blinded assessors, subjects not blinded to treatment but with an effort not to bias them towards any particular exercise treatment, protocols provided, a large study loss of 30% with intention to treat analyses performed, and power analyses to determine sample sizes. The PEDro score for the article by Koumantakis *et. al* (2005) was 6/10 and was 7/10 for the article by Cairns *et. al* (2006), with the difference in PEDro scores because VAS B scores between groups that were not similar at baseline in the article by Koumantakis *et. al* (2005), which means that randomization may not have been completely successful, but this was taken into account by the authors. Overall, both articles closely matched my clinical question. Both studies compared specific local core stabilization exercises to conventional PT treatment including global core stabilization exercises and used the RMDQ as a measure of functional outcome and the VAS score as the pain measurement scale. I used the outcomes from both studies to answer my clinical question, with slightly greater weight placed on the study by Koumantakis *et. al* (2005) even though the PEDro score was one lower because it most closely matched my clinical PICO. The detailed protocol for the comparison group in the study by Koumantakis *et. al* (2005) showed that global core stabilization exercises were included, but it is difficult to determine what kind of global core stabilization was included in the comparison group with conventional PT in the study by Cairns *et. al* (2006) without the detailed protocol.

The subject populations in both studies were similar to my clinical population in terms of inclusion and exclusion criteria, with the exception that younger subjects were included in both studies and the study by Cairns *et. al* (2006) included subjects from UK. It is possible that there were some cultural differences affecting the rating of function on the RMDQ and pain on the VAS, which may have contributed to the lack of statistical significance on the RMDQ and VAS scores between groups. Overall, I would be comfortable generalizing the results of these 2 studies to a population including patients aged

18 - 65 years with chronic LBP not due to instability for at least 6 months affecting activities of daily living (ADLs), mobility, work or recreational activities.

In both studies, the comparison was specific core stabilization with conventional PT including general exercise, global core stabilization and pain education to conventional PT, and did not include a comparison to no treatment. Clinically, one would decide between the addition of specific core stabilization first to conventional PT and then progressing to global core stabilization and incorporating into functional activities, so both studies address this question well. However, since the detailed protocol for the comparison group in the study by Cairns *et. al* (2006) was not provided, it is unclear as to whether global core stabilization was included in conventional PT in that study, but is relatively safe to assume that at least some degree of global core stabilization was included based on knowledge of what is common clinical practice for patients with chronic LBP.

The treatments in both studies were feasible and easily provided at any clinic. In the study by Koumantakis *et. al* (2005), a total of 16 sessions was administered at 2 times a week for 8 weeks, each 45-60 minutes long. There were fewer and shorter sessions administered in the study by Cairns *et. al* (2006), with up to a maximum of 12 sessions over 12 weeks, each 30 minutes long. Both time frames are within reason to be covered by insurance and feasible to administer.

Both studies found a statistically significant and clinically important improvement in function and pain within the intervention and comparison groups, but not between groups at the follow up which was 3 months (Koumantakis *et. al*, 2005) to 12 months (Cairns *et. al*, 2006) after discharge. The results are summarized on Table 9. Thus, there is no conclusive evidence to suggest that including specific core stabilization is any more effective than conventional PT including global core stabilization in improving long term function and pain in patients with chronic LBP. Based on the results of these 2 studies, I would recommend that the clinician use discretion in determining whether to include specific segmental core stabilization in treatment of patients with nonspecific chronic LBP on a case by case basis.

Table 9: Summary of RMDQ and VAS mean differences with 95% CI

	TG (Specific stabilization)		CG (Global stabilization)		Between groups	
	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)	RMDQ (0-24)	VAS (10 cm)
Baseline to discharge (8 wks)[1]	-4.05#* (-7.31, -0.79)	-18.18#* (-36.98, 0.62)	-6.60#* (-11.57,-1.63)	-14.92#* (-31.44, 1.60)	2.55 (0.30, 4.81)#*	-3.26 (-10.15, 3.63)
Baseline to 3 month follow up[1]	-4.65#* (-7.91, -1.39)	-15.16#* (-34.26, 3.94)	-6.03#* (-11.01,-1.05)	-17.78#* (-37.48, 1.92)	1.38 (-0.87, 3.64)	2.62 (-4.58, 9.82)
Baseline to 12 month follow up[2]	-4.5#* (-6.2, -2.9)	-1.2#* (-2.1, -0.4)	-5.2#+ (-6.7, -3.6)	-1.8#* (-2.6, -0.9)	0.6 (-2.9, -1.7)	-0.5 (-1.75, 0.72)

# indicates a statistically significant difference was found
+ indicates the mean difference and 95% CI exceeded MCID
* indicates the mean difference (but not 95% CI) exceeded MCID
1 Study by Koumantakis *et. al* (2005)
2 Study by Cairns *et. al* (2006)

Recommendations for future research: Overall, both studies included a population similar to the clinical population of interest, had blinding of assessors, power analysis to determine sample size, addressed reliability, validity and MCIDs for the RMDQ and VAS scores, and described the specific core stabilization protocols well. However, there were large study losses with intention to treat analyses to account for that sampling from one center in the study by Koumantakis *et. al* (2005), lack of detailed protocol and uneven number of treatment sessions between intervention and comparison groups in the study by Cairns *et. al* (2006). In future studies, I would recommend sampling from more than one facility, trying to minimize study losses or using a larger sample size, including a more detailed protocol for the comparison group as well and using similar number of treatment sessions and time for both intervention and comparison groups. Also, it might be beneficial to take a look at recurrence rate of LBP episodes as one of the outcome measures to determine if specific segmental core stabilization is more effective than global core stabilization in the prevention of recurring LBP episodes.

References:

1. Critchley DJ, Ratcliffe J, Noonan S, Jones RH, Hurley MV. Effectiveness and cost-effectiveness of three types of physiotherapy used to reduce chronic low back pain disability: a pragmatic randomized trial with economic evaluation. *Spine* 2007;32(14):1474-81.
2. Koumantakis GA, Watson PJ, Oldham JA. Trunk muscle stabilization training plus general exercise versus general exercise only: randomized controlled trial of patients with recurrent low back pain. *Physical Therapy* 2005;85(3):209-25.
3. Cairns MC, Foster NE, Wright C. Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine* 2006;31(19):E670-81.
4. Rackwitz B, de Bie R, Limm H, von Garnier K, Ewert T, Stucki G. Segmental stabilizing exercises and low back pain. What is the evidence? A systematic review of randomized controlled trials. *Clinical Rehabilitation* 2006;20(7):553-67.
5. Kumar S, Sharma VP, Negi MP. Efficacy of dynamic muscular stabilization techniques (DMST) over conventional techniques in rehabilitation of chronic low back pain. *Journal of Strength and Conditioning Research* 2009;23(9):2651-9.
6. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 1996;65(1):71-6.
7. Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study Group. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *European Spine Journal* 2003;12(1):12-20.
8. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. Low back pain: outcome measures. *Journal of Rheumatology* 2001;28(2):431-8.
9. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskeletal Disorders* 2006;7:82-97.
10. Ostelo RW, deVet HC. Clinically important outcomes in low back pain. *Best Practice & Research in Clinical Rheumatology* 2005; 19(4):593-607.
11. BlueCross BlueShield Association Technology Evaluation Center. Special Report: Measuring and reporting pain outcomes in randomized controlled trials. *Assessment Program* 2006;21(11). Accessed at http://www.bcbs.com/blueresources/tec/vols/21/21_11.html on February 6, 2011.